DOCUMENT RESUME

ED 114 409 TM 004 892

AUTHOR Anderson, Lorin W.
TITLE Opportunity to Learn, Test Bias, and School Effects.
PUB DATE 75
NOTE 18p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Washington, D.C., March 30-April 2, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Academic Achievement; *Achievement Tests; Culture Free Tests; Elementary Education; *Incidental Learning; Standardized Tests; *Test Bias; Test Validity; *Vocabulary

ABSTRACT

The purpose of the study was to examine test bias and the "non-effects" of schooling. Teachers were given a list of words selected from standardized vocabulary tests and asked to indicate the words they had taught. The words were classified by the grade level at which they were first introduced. Ninety-five third-grade students in four schools were given two 17-item subtests: subtests of items the students did and did not have an opportunity to learn. The difference in the scores on the tests was used as an index of school effects of student vocabulary. Achievement test bias (when equated with content validity) was present. Students scored higher on the subtest containing items they had had an opportunity to learn. (Author/RC)

**********
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
**********

ED114409

# Opportunity to Learn, Test Bias, and School Effects

Lorin W. Anderson

University of South Carolina

Paper presented at the Annual Meeting of the National Council of Measurement in Education, Washington, March 30 - April 2, 1975.

TM004 892

## Abstract

The purpose of the study was to examine test bias and the "non-effects" of schooling. Teachers were given a list of words selected from standardized vocabulary tests and asked to indicate the words they had taught. The words were classified by the grade level at which they were first introduced. Ninety-five third grade students in four schools were given two seventeen item subtests: subtests of items the students did and did not have an opportunity to learn (Test Y and Test N, respectively). The difference in the scores on Test Y and Test N was used as an index of school effects on student vocabulary. Achievement test bias (when equated with content validity) was present. Students scored higher on the subtest containing items they had had an opportunity to learn ($p < .01$).

Opportunity to Learn, Test Bias, and School Effects

## Introduction

Two conclusions which seem to have been largely accepted by educators in recent years concern the test bias of various standardized tests (Thorndike, 1971) and the apparent non-effects of schooling (Coleman, 1966). The acceptance of research findings concerning the bias of tests for students living in minority subcultures was largely responsible for the attempted movement away from the use of tests altogether or, at the very least, the movement toward "culture fair" tests. The acceptance of the findings concerning the impotence of schools in producing student learning was in part responsible for the anti-school or de-schooling movement. This paper is an attempt to examine the questions of test bias and non-effects of schooling in a different light.

Recently, Lewy (1972) investigated the importance of a variable termed "opportunity to learn" in relationship to student achievement. Opportunity to learn was defined as the extent to which the teachers believed that their students had had an opportunity to learn each item on a particular achievement test. Lewy concluded from his research that a relationship did exist between opportunity to learn a topic and proficiency in that topic. In view of these findings it seems as though the variable "opportunity to learn" can be useful in examining both test bias and the possible effects of schooling on student achievement.

The term "culturally biased" has most frequently been applied to tests of mental ability and aptitude. It seems as though the term would be applied very differently to achievement tests. Whereas intelligence tests can be termed "biased" to the extent that certain items do, in fact, favor one subcultural group over another, achievement tests can only be biased to the extent that the achievement tests do not measure what is taught in the schools (i.e., to the extent that they are not content valid). In other words, if students have not had an opportunity to learn particular items in their school, those particular items are "biased" against all students in that school.

If "biased" and non-biased items can be identified in this manner, then it may be possible to assess the effects of schooling. Schooling would be said to have an effect on student learning to the extent that students perform better on a test containing items that they have had an opportunity to learn (unbiased items) than on a test containing items that they have not had an opportunity to learn ("biased" items).

In line with the above discussion the following three null hypotheses were investigated in the present study:

1. There will be no racial nor socio-economic status differences in the number of vocabulary words that students have had an opportunity to learn. In other words, schools of various socio-economic status and racial make-up will not differ in the number of vocabulary words they teach their students.

2. There will be no racial or socio-economic status differences on a subtest containing words that students did have an opportunity to learn in school (Subtest Y) nor on a subtest of words that the students did not have an opportunity to learn in school (Subtest N).

3. There will be no difference between the mean scores on Subtest Y and Subtest N across all students.

Procedures

Three standardized tests were used to select a pool of items. All of the items in the pool were chosen from the Reading Vocabulary Test, Primary Battery, Grades 2 through 4. The number of vocabulary words selected was 114. Slight modifications in the item format were made when necessary to insure similarity of format for all of the items.

A list of the vocabulary words, excluding the item response alternatives, was given to all first through fourth grade teachers in four elementary schools in a metropolitan area in South Carolina.

The directions which accompanied the vocabulary list were as follows: "Place a check in front of each word on the accompanying list that you have taught in a vocabulary lesson or incidentally as a vocabulary word. 'Incidentally as a vocabulary word' means that you took time to instruct the class in the meaning of a new word when they encountered it in their course readings."

On the basis of the results of this survey each word was classified at the particular grade level at which it was first taught. A word was classified at a particular grade level if two-thirds or more of the grade-level teachers indicated that they had taught the word. A table was then drawn up for each school indicating the grade level placement of each word. From this table, two sub-lists of words were selected. The first list contained the words that were taught in all the schools in either the second or third grade ($i = 17$). These were designated as the words the students had had an opportunity to learn (Test Y). The second list contained the words that were

either taught in the fourth grade in all schools or not taught at all in grades one through four (i = 17). These were designated as the words that none of the students in any of the four schools had had an opportunity to learn (Test N). Words taught in grade one were eliminated because it was felt that they were much less difficult words.

The thirty-four items were then assembled into a single test and administered to all of the third grade students in four schools. The alternatives used for each item were the same as those used in the standardized test from which the item was taken. KR20 reliability estimates for the two subtests were .66 for Test Y and .67 for Test N.

Sample

The four schools were selected because of their racial and socio-economic status composition. School I was a middle class, virtually all-white school. School II was a middle class, racially mixed school. School III was a lower class, racially mixed school. School IV was a lower class, virtually all black school. The socio-economic status designation for each school was based on the median father's occupational level and the median parents' educational level taken from the school records.

The sample used in the study was chosen in the following manner. All of the black third grade students in School II (n = 21) and all of the white third grade students in School III (n = 24) were selected. Twenty-five white third grade students were randomly selected from School I and twenty-five black third grade students were randomly selected from School IV.

Results

Table 1 indicates the number of words introduced at each grade level in each of the four schools.

---

Insert Table 1 About Here

---

The first grade students in the white, middle class school have been exposed to twice as many words as the black, middle class students, and approximately five times as many words as students in the two lower class schools. After this initial "head start" the number of words introduced per year does not differ a great deal across the schools. However, when this head start is examined cumulatively the school differences in opportunity to learn appear large.

In order to further examine this cumulative difference a Verbal Knowledge Index (VKI) was calculated for the third grade students in the four schools. The VKI is computed by multiplying the number of words introduced each year by the number of years a third grade student would have had access to that word and summing up the products. An example of the computation of the index is given in the note in Table 2.

---

Insert Table 2 About Here

---

If we interpret the VKI as a rough index of the students' "school-induced" working vocabulary, the results indicate that this working vocabulary of the third grade students in the middle class schools is twice as large as that of the lower class third graders. Also the working vocabulary of the white third grade students is about one and one-fourth times as large as the black third graders.

These data suggest that the original list of vocabulary words is not equally content valid for the students in the four schools. The list seems to favor the middle class students over the lower class, and, to a lesser extent, the white students over the black. This "favoritism" indicates differences in schools and schooling rather than in home background.

In order to investigate the last two hypotheses two subtests were formed. One consisted of words that were said to have been taught in grades two or three in all schools. The other contained words that were not taught in grades one, two, or three in any of the schools. Table 3 contains the words that comprise the two subtests.

---
Insert Table 3 About Here
---

An index of "school effect" was computed by subtracting each student's score on Test N (the non-taught words) from the score on Test Y (the taught words). A positive score on this index means that the student scored higher on the words which he had had an opportunity to learn. A negative score means that the student scored higher on the words which he had not had an opportunity to learn. Further, a zero mean score for a school indicates that the scores on the two subtests are the same (i.e., schooling has no effect).

A two-way analysis of variance (Race x SES) was used to investigate the second hypothesis. The results of the analysis of variance are found in Table 4.

---
Insert Table 4 About Here
---

The Race and SES main effects are not significant. School effects are not significantly different for black or white students, nor for middle and lower socio-economic status students. The interaction term (Race x SES)

approaches statistical significance. A graphical display of this interaction is shown in Figure 1.

---
Insert Figure 1 About Here
---

In examining Figure 1 it appears that the school effects are virtually the same for the white middle and white lower class students. On the other hand the school effect appears to be much more marked for the black middle class students than the black lower class students. The school effect is significantly different from zero for the black middle class students ($p<.01$), and for the middle and lower class white students ($p<.05$). The school effect is not significant for the lower class black students.

This data tends to suggest that the black middle class students benefit the most from schooling in terms of their vocabulary. On the other hand, black lower class students benefit very little, if any.

Across all students the schooling effect is significantly positive. Table 5 shows the means, standard deviations, and correlated t-test for all students on Test Y and Test N.

---
Insert Table 5 About Here
---

Students scored significantly higher on Test Y than on Test N. However, despite this significant difference, the mean scores for Test Y do not seem to be very high (approximately eight correct out of seventeen items).

On the basis of the above observation the teachers were questioned after the results of the study had been examined. The teachers were asked to list the synonyms they used to teach their students about each of the words in Test Y. The actually taught synonymns were then compared with alternatives available for each item. The average number of synonyms that, in fact,

appeared as correct choices to the items on Test Y was slightly less than fifty per cent. In other words, even though the teachers had, in fact, taught the words, in many cases they had used different synonyms and/or explanations in class than those which appeared as item choices.

Conclusions and Implications

It has been suggested that when achievement tests are used, test bias be equated with content validity. When this is done, the results of the study suggest that the vocabulary subtests of various standardized tests are biased in favor of white and middle class students. The social class bias appears to be stronger than the racial bias.

One potential problem in establishing content validity concerns the "correct" answer called for by the test manufacturer. In order to establish content validity it seems necessary to determine the answers to two separate, but related,questions. First, did the student have an opportunity to learn this content? Second, did the student learn the content in a manner which will allow him to answer the item correctly?

It has further been suggested that the discrepancy of the scores on a test containing items the students have had an opportunity to learn and a test containing items they have not had an opportunity to learn be used as an index of school effects. The results of this study suggest that when this index is used, schools do have a significant effect on achievement. More specifically, it is concluded that schooling is most important for the development of the vocabulary of black, middle class students. Schooling seems to be virtually ineffective in developing the vocabulary of black, lower class students.

## References

Coleman, J.S. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Lewy, A. Opportunity to Learn and Achievement in Three Subject Matter Areas. Journal of Experimental Education, 41, 1972, 68-73.

Thorndike, E.L. Concepts of Culture-Fairness. Journal of Educational Measurement, 8, 1971, 63-70.

Table 1

Number of Words Introduced at Each Grade Level

| Type of School | Grade 1 | Grade 2 | Grade 3 | Grade 4 | Not Taught |
|---|---|---|---|---|---|
| White, Middle | 71 | 7 | 16 | 15 | 5 |
| Black, Middle | 32 | 31 | 22 | 17 | 12 |
| White, Lower | 15 | 15 | 34 | 20 | 30 |
| Black, Lower | 12 | 13 | 32 | 23 | 34 |

Table 2

Verbal Knowledge Index for Grade 3 Students in Each School

| Type of School | Verbal Knowledge Index | Ratio of VKI(School X/ Black, Lower) |
|---|---|---|
| White, Middle | 243 | 2.58 |
| Black, Middle | 180 | 1.91 |
| White, Lower | 109 | 1.15 |
| Black, Lower | 94 | 1.00 |

Note. - The Verbal Knowledge Index is computed by multiplying the number of words introduced each year by the number of years a third-grade student would have had access to that word. For example, a third grade student in the white, middle class school was introduced to 71 words in first grade. By the third grade he would have had an opportunity to use that word for three years. Thus for his first grade year the verbal knowledge index is 71 x 3 or 213. For the second grade the VKI is 7 x 2 or 14 and for the third grade the VKI is 16 x 1 or 16. The composite index is the sum of these products, i.e., 213 + 14 + 16 = 243.

Table 3

List of Words Taught and Non-Taught in School

| Test Y (Taught in Grades 2 or 3) | Test N (Not Taught in Grades 2 or 3) |
|---|---|
| Perhaps | Pretending |
| Slip | Select |
| Ideas | Examine |
| Sure | Journey |
| Decide | Manufactured |
| Village | Precious |
| Smiling | Refinements |
| Instrument | Avenue |
| Path | Shiver |
| Hopped | Approach |
| Carry | Destroy |
| Investigate | Halt |
| Discover | Platform |
| Bald | Offer |
| Unusual | Deliver |
| Envelope | Defeat |
| Multiply | Model |

Note. -- Non-Taught in Grades 2 or 3 means that the word was taught no earlier than Grade 4.

Table 4

Analysis of Variance of the Effect of Schooling on Racial and Socio-Economic Status Groups

| Source | df | SS | MS | F-value | p |
|---|---|---|---|---|---|
| Race | 1 | 0.91 | 0.91 | 0.15 | .70 |
| SES | 1 | 9.65 | 9.65 | 1.64 | .20 |
| Race x SES | 1 | 16.32 | 16.32 | 2.78 | .10 |

Table 5

Means, Standard Deviations and Correlated t-test for the Total Sample on those Words that Were and Were Not Taught in School

(n=95)

| Subtest | Mean | Standard Deviation | t | p |
|---|---|---|---|---|
| Taught Words (Test Y) | 7.72 | 3.15 | 3.60 | .01 |
| Non-Taught Words (Test N) | 6.82 | 3.05 | | |

Figure 1

A Schematic Diagram of the Differential Effects of Schooling on Various Racial, Socio-Economic Status Students